

AR-010-173

DSTO-TR-0511

O

F

S

D

"Merge" – A Filter for the Fusion of
Dual-Frequency Sidescan Sonar Data

Roger A. Neill and Stuart D. Anstee

] APPROVED FOR PUBLIC RELEASE

© Commonwealth of Australia

19970724 083

DEPARTMENT OF DEFENCE
DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

"Merge" - A Filter for the Fusion of Dual-Frequency Sidescan Sonar Data

Roger A. Neill and Stuart D. Anstee

**Maritime Operations Division
Aeronautical and Maritime Research Laboratory**

DSTO-TR-0511

ABSTRACT

A filtering and data fusion technique is described which uses the correlation between the two data streams of a dual-frequency sidescan sonar in order to discriminate against noise and preferentially enhance features of desired relative size. The filter is applied to data collected with a Klein 590 dual-frequency sidescan sonar with the parameters adjusted so as to enhance the display of minelike objects. Experience with the filter shows that it is highly effective in noisy conditions.

DTIC QUALITY INSPECTED 2

RELEASE LIMITATION

Approved for public release

D E P A R T M E N T O F D E F E N C E

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

Published by

*DSTO Aeronautical and Maritime Research Laboratory
PO Box 4331
Melbourne Victoria 3001*

Telephone: (03) 9626 7000

Fax: (03) 9626 7999

© Commonwealth of Australia 1997

AR-010-173

March 1997

APPROVED FOR PUBLIC RELEASE

"Merge" - A Filter for the Fusion of Dual-Frequency Sidescan Sonar Data

Executive Summary

The Klein 590 sidescan sonar which is incorporated in the Royal Australian Navy (RAN) Route Surveillance System (RSS) simultaneously returns two streams of data (sonograms), with identical data rates, corresponding to two sonar operating frequencies, 100 kHz and 500 kHz. Each sonogram corresponds to a two-dimensional acoustic "map" of the seabed and corresponding pixels in each map correspond to areas on the seabed with the same centre, but generally not the same size. The beams of the sonar also have different characteristics for the two operating frequencies and separate electronic circuits to generate and receive them. There is a high degree of correlation between the seafloor "signal" component of the two data streams, but noise inputs tend to be less well correlated between the two frequencies. This report details a method of "data fusion" based on a modification to the correlation coefficient expression for a linear regression, which when applied to the two sonograms, results in a third, "merged" sonogram (and corresponding "map" of the sea floor) which discriminates against noise in favour of signals from objects of interest. In particular, the technique may be adjusted so as to discriminate in favour of objects of minelike sizes and to reject bottom "clutter", surface reverberation noise and other sources of noise such as electrical interference.

In this report, the algorithm used to produce merged sonograms is discussed. A modification of the correlation coefficient for a linear regression is shown to be superior to the correlation coefficient in the determination of "correlation" between the two original sonograms and some suggestions are given as to why such a technique should be successful. The implementation of the merging algorithm as a real-time process suitable for routine use during data collection or post-processing is also discussed. Some example images are shown which give an indication of the components of the process and the usefulness of the algorithm. Following the discussion of the algorithm, an attempt is made to quantify the usefulness of the algorithm by the definition of a "local signal to noise ratio" which gives a measure of how much a given minelike object can be expected to stand out from the background image clutter. Such a local signal to noise ratio has been evaluated for hundreds of minelike objects, and a comparison is made of the values recorded from both original sonograms (100 and 500 kHz) and from the merged sonogram. The values from the merged sonogram are shown to be consistently superior to those from the original sonogram, usually by factors of the order of 6 dB for the median scores. The superiority of the appearance of minelike objects in the merged sonograms applies regardless of range from the sonar and weather conditions.

In summary, the merging algorithm described in this report appears to be robust and useful in routine operations where dual-frequency sidescan sonar data is available and the location of minelike objects is of interest.

Authors

Roger A. Neill Maritime Operations Division



Roger Neill received a Science Degree with Physics majors in 1978 and a PhD in 1984. The subject of his doctoral dissertation was the development of new methods for evaluating human binocular visual function. He then conducted post doctoral research on this project until he joined DSTO in 1988. Upon joining DSTO Roger initially investigated the effects of flicker and clutter on target detection in screen-based equipment. He subsequently worked towards the development of improved methods for the operational use of the Royal Australian Navy's sidescan sonar-based Route Surveillance System, and is currently investigating the application of new-technology sidescan sonars to route surveillance.

Stuart D. Anstee Maritime Operations Division

Stuart Anstee is a Research Scientist in the Sydney branch of MOD. He holds a BSc in Applied Mathematics and Physics, an Honours degree in Physics and a PhD degree in Physics from the University of Queensland. His research concerns underwater acoustics, image processing and sonar design issues related to mine warfare.

Contents

1. INTRODUCTION	1
2. FILTER IMPLEMENTATION	2
2.1 Determination of correlation between sonograms	2
2.1.1 Use of the correlation coefficient for a linear regression	3
2.1.2 A modified correlation coefficient	4
2.1.3 Making a useful filter	6
2.2 Filter implementation	8
2.3 Some examples	9
2.4 Performance evaluation	14
2.4.1 Definition of "local signal to noise ratio" for a target object in a sonogram....	14
2.4.2 Variability of the local signal to noise ratio	15
3. FILTER OPERATION AND PERFORMANCE	16
3.1 Analysis of Jervis Bay data sets up to 1995 with a fixed filter parameter.....	17
3.2 Analysis of Jervis Bay 1996 data sets collected with different towfish configurations	20
3.3 Why the filter works	22
4. CONCLUSION.....	22

1. Introduction

The majority of sidescan sonars used for the Route Surveillance System (RSS) of the Royal Australian Navy (RAN) are a dual-frequency device, the Klein System 590, which simultaneously outputs sonogram data for two operating frequencies, 100 kHz and 390 kHz, known however as "500 kHz" in the sonar industry and hereafter. The two data streams are highly correlated in space and time - the beams are emitted from parallel transducer elements separated by less than 25 millimetres in space and by 60 μ s in time, far shorter than significant transducer motion time-scales. The beams emitted by the transducers share a common central plane and only differ in their spreading characteristics, and in their attenuation with range. Because of the high degree of correlation between return signals from the two beams inherent in this sonar system, it is possible to discriminate against noise by combining the data from the two channels to form a single sonogram. It is often an observable feature of dual-frequency sonograms that the noise characteristics of the two frequencies are quite different - one frequency may, for instance, be subject to surface reverberation while the other may be degraded by electrical interference. If this "merged" sonogram is of sufficient quality, it should be sufficient to replace the two original sonograms and provide the default interface during processing of route-survey or sidescan-based minehunting data, or, at worst, it should provide a useful adjunct to the images associated with the individual sonar frequencies.

To date, only a single implementation of a "merging" process is known to the authors. The "Data Fusion" © process implemented by EG&G Marine Instruments Corporation in their DF-1000 digital dual-frequency (100-500 kHz) sidescan sonar system is directed at making the best use of the range and resolution characteristics of each sonar frequency. The sonar operator sets a range control, and for slant ranges below the control value, the higher-resolution 500 kHz data is displayed. At higher ranges, the proportion of the 100 kHz data is increased to counter the attenuation of the 500 kHz data and for ranges greater than a second range value, only 100 kHz data is displayed. This process, which may be of value in itself, does not attempt to use the correlation between the two sonograms in order to improve the image.

In this report, a merging algorithm which appears to be sufficiently robust for use in routine route-survey data inspection, and sufficiently simple to be implemented as a real-time process operating in the background of a display or acquisition program, is discussed. The filter was developed empirically and represents the best of many attempts. Optimal filter parameters have been determined empirically and these are also discussed. A numerical measure of the effectiveness of the filter in revealing the presence of minelike objects against a noisy background is developed and the performance of the filter is assessed using the measure applied to filtered output from data collected in sea-trials in Jervis Bay.

2. Filter implementation

2.1 Determination of correlation between sonograms

Consider two sonograms U and V of the same scene, collected simultaneously with beams of different frequency. Assume that the sonograms consist of discrete arrays of samples $u[h,k]$ and $v[h,k]$ where h corresponds to the h th pulse recorded and the h th row of the sonogram displayed as an image, and $k=0,1,\dots,N-1$ describes the k th time bin and corresponds to the k th column of the image. Each ordered pair $[h,k]$ labels a particular location in both sonograms, and labels two physical regions on the sea floor which have the *same* centre, although their size will in general be different because of the different widths of the sonar beams. The sampling regime does not vary from pulse to pulse and is the same for both frequencies, so that h and k may be treated as independent variables.

Let S be a subset of M locations in the sonogram, say $S \equiv \{[h,k]_i, i = 0,1,\dots,M-1\}$. The shape of the region chosen by S is arbitrary, but it could, for example, describe a closed area on the sea bed, say a rectangular or circular patch.

Note that there are some assumptions inherent in the above definitions. As noted, h and k are independent variables, since the returns from all pulses are sampled the same way. In addition, at the frequencies of interest, the seawater medium may be taken to be linear and non-dispersive, which means that for each pulse, the same array of sampling time bins can be used for both frequencies. As noted in the introduction, the pulse at one frequency is generated before the other one and the sampling times for the second frequency are delayed correspondingly. In the absence of platform motions, therefore, each pair of samples $[h,k]_i$ corresponds to a pair of physical areas on the sea bed with the same centre. In practice, the interval between the pulse generations is small enough in comparison to the time-scales for sonar platform motion, that beam motion may be neglected.

The filtration described in this report depends on two properties, neither one being sufficient by itself; when the region S contains an object of interest or a shadow:

- a) The absolute value of the divergence of the signal level in S from the background level must be large in at least one of the sonograms; and
- b) There must be a correlation between the signal levels $u[h,k]_i$ and $v[h,k]_i$ in S , such that when condition "a" is true, then the divergence of the signal level in the *other* sonogram away from the background level is also non-negligible.

A necessary, additional condition placed upon the filter was that it must be able to perform satisfactorily when applied to sonar data which contain significant across-track variations in background signal level. Side scan sonars use time-varying gain

(TVG) algorithms to correct for such across-track effects, but it is not unusual to find residual variability due to slight mismatches between the sonar's TVG algorithm and the actual across-track signal structure.

2.1.1 Use of the correlation coefficient for a linear regression

A method is now required in order to discriminate between correlated and uncorrelated regions of the sonograms. Let the two data sets in the region S be abbreviated by $x_i \equiv u[h, k]_i$ and $y_i \equiv v[h, k]_i$. Both have the character of positive semi-definite random variables. If the region S is mostly occupied by an object of interest or a shadow, then property "b" above implies that there should be a strong relationship between the two data sets. The simplest relationship is a linear one of the form

$$y_i \approx m x_i + c \quad (2.1)$$

where m and c will depend on factors like the gain adjustments of the sonar. This is simply saying that an increase in the relative signal level in one sonogram due to a scatterer is likely to be accompanied by a proportional increase in the other sonogram. The fundamental measure of linear correlation between the data sets is given by the raw cross-correlation

$$\langle xy \rangle = \frac{1}{M} \sum_{i=0}^{M-1} x_i y_i \quad (2.2)$$

which, since x_i and y_i are random variables, should peak when they are well-correlated. However, by itself, the cross-correlation is not a useful measure, because it is not normalised. It is better to use the correlation coefficient for a linear regression, r , which is defined by

$$r = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\left[(\langle x^2 \rangle - \langle x \rangle^2) (\langle y^2 \rangle - \langle y \rangle^2) \right]^{1/2}} \quad (2.3)$$

where the notation $\langle f \rangle$ refers to the mean value

$$\langle f \rangle = \frac{1}{M} \sum_{i=0}^{M-1} f_i \quad (2.4)$$

The correlation coefficient r will have a value between -1 and 1. The value $r = +1$ occurs when x_i and y_i are perfectly linearly correlated, that is, when the relationship (2.1) applies for any positive value of m and c ; $r = 0$ occurs when they are completely uncorrelated; and $r = -1$ occurs when they are anti-correlated, that is when m is negative. Negative correlation coefficients are unlikely to correspond to objects of interest, since the signals from the two frequencies should both rise on encountering a strong reflector and both fall in a shadow zone, although generally to different extents.

As it stands, the correlation coefficient is not particularly useful, since it is found that neither minelike objects nor strong shadows tend to correspond to the most strongly correlated regions in the sonograms. If the target region contains strong shadow, then the signal level is minimal and is dominated by noise which is not correlated between the two data sets, so the value of r is small. In the case of a strong reflector, two factors work against large values of the correlation coefficient. Firstly, constructive and destructive interference induces a complicated reflected sound field which varies greatly over small angular separations. In this case, the variation in return from a reflective object has the character of multiplicative noise - this tends to increase the standard deviations σ_x and σ_y in the neighbourhood of extended reflective objects. Since the sound-field is frequency-dependent, the correlation between the data sets is reduced. Secondly, the dynamic range between the background and strong acoustic reflectors is so large that the signals from strong reflectors tend to saturate or "clip". If the target region is clipped in both data sets, then the correlation may increase, however if only one is clipped, which often occurs, the correlation decreases.

What is required is a measure which takes account of the property "a" mentioned previously, that minelike objects tend to have quite different mean signal levels than their surroundings.

2.1.2 A modified correlation coefficient

It was discovered by experiment that a modification of the correlation coefficient expression (2.3) gives a more useful coefficient than the exact expression. Expression (2.3) can be re-arranged as

$$r = \frac{\langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle}{\left[\langle (x - \langle x \rangle)^2 \rangle \langle (y - \langle y \rangle)^2 \rangle \right]^{1/2}} \equiv \frac{\langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle}{\sigma_x \cdot \sigma_y} \quad (2.5)$$

which makes more obvious the fact that values other than the true means $\langle x \rangle$ and $\langle y \rangle$ can be substituted into the expression, that is, a pseudo-correlation coefficient can be defined as

$$r' = \frac{\langle (x - \bar{X})(y - \bar{Y}) \rangle}{\left[\langle (x - \bar{X})^2 \rangle \langle (y - \bar{Y})^2 \rangle \right]^{1/2}} \quad (2.6)$$

where \bar{X} and \bar{Y} have been substituted for the means of the data sets x_i and y_i . Schwarz's inequality shows that r' is still limited to the range $[-1,1]$, whatever the values of \bar{X} and \bar{Y} . The pseudo-correlation coefficient becomes useful when \bar{X} and \bar{Y} are taken to mean the local background levels near S in each of the two sonograms. Thus, consider a region S' which is larger than the target region S and is in some way local to S . The shape of S' is not important, except that it constitutes a "background"

against which the smaller set S can be compared. Let the points in S' be known as X_i and Y_i , with means \bar{X} and \bar{Y} respectively. Now, let the means of x_i and y_i in S be referred to as the "target averages" and means of X_i and Y_i be referred to as the "background averages". Define the differences between the target and background averages (the "mean offsets") in each sonogram by

$$\begin{aligned}\alpha &= \langle x \rangle - \bar{X} \\ \beta &= \langle y \rangle - \bar{Y}\end{aligned}\tag{2.7}$$

then substituting for \bar{X} and \bar{Y} in (2.6) and simplifying, we have

$$r' = \frac{r\sigma_x\sigma_y + \alpha\beta}{\left[(\sigma_x^2 + \alpha^2)(\sigma_y^2 + \beta^2)\right]^{1/2}}\tag{2.8}$$

Inspecting expression (2.8), we can see that it can be written in terms of two "angles", say ξ_x and ξ_y , which show the relationships between the respective target standard deviations and the offsets of the target means from the background means.

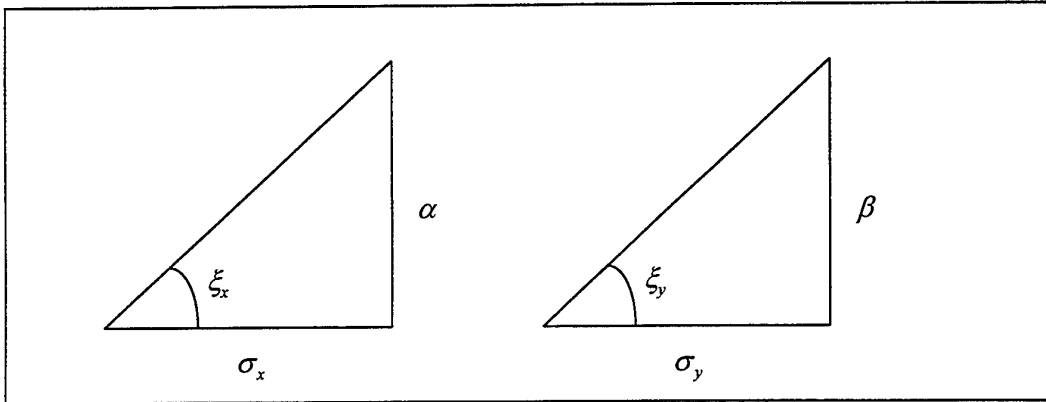


Figure 1. Relationships between the "angles" ξ_x and ξ_y , target standard deviations σ_x and σ_y and mean offsets α and β .

Figure 1 shows the significance of the angles, which allow expression (2.8) to be written as

$$r' = r \cos \xi_x \cos \xi_y + \sin \xi_x \sin \xi_y\tag{2.9}$$

where r' , r and the sinusoidal quantities all lie in the interval $[-1, 1]$. Some observations can be made about the mapping:

- a) The target standard deviations σ_x and σ_y are positive semi-definite, so that the cosines of each angle must also be, that is, the angles ξ_x and ξ_y must lie in the range $[-\pi/2, \pi/2]$. Thus, increasing pixel-by-pixel correlation r in the target region always increases the value of r' .
- b) When ξ_x and ξ_y are both of small magnitude, then the cosine terms will be close to +1 and the sine terms will be small. In this case, r' will be close to r .
- c) When ξ_x and ξ_y both have magnitudes close to $\pi/2$, then the cosine terms become small and the sine terms dominate. In this circumstance, the value of r is of negligible importance.
- d) When only one of ξ_x and ξ_y has a magnitude close to $\pi/2$, then the value of r is likely to be important to the value of r' .

Considering the points with regard to the use of the operator in highlighting objects of interest:

- i) Point "a" implies that the coefficient is well-behaved with regard to correlation in the target region, since r' increases with increasing correlation. This is a necessary feature.
- ii) Point "b" applies when α and β are small in comparison to σ_x and σ_y respectively, that is, when the target mean offsets from the background are smaller than the standard deviations in the target regions. This will apply when the target region contains the normal "background", in which case, the magnitude of r' will depend on the correlation r inside the target region. In general, the correlation between the data sets increases when the target region contains a strong reflector, which is not the case for background regions. Thus, the coefficient tends to suppress background regions.
- iii) Point "c" applies when α and β are large in comparison to σ_x and σ_y respectively, that is, when the target mean offsets from the background are larger than the standard deviations in the target regions. In this case, as noted, r' can be large regardless of whether there is significant correlation between the two target regions. Thus, the coefficient can enhance both strong reflectors and strong shadow regions.

2.1.3 Making a useful filter

The criteria given in the preceding section make it clear that if the target regions S and S' are chosen correctly, then the pseudo-correlation coefficient should discriminate in favour of target regions with strong correlation between the data sets and large offsets

between the target and background mean values. In order to maximise the likely offset between background and target mean values, the background region should be chosen large enough for variations in the target region to have negligible impact on the background mean, but small enough to respond to changes on scales much larger than likely objects of interest. Similarly, the target region should be chosen to be as compact as possible in order to maximise resolution, but sufficiently large to encompass a significant portion of the typical target size and establish meaningful values for the means and standard deviations.

Due to the nature of sonar signals, the pseudo-correlation coefficient alone is not sufficient to yield a useful sonogram because of two limitations. Firstly, the resolution of the pseudo-correlation coefficient is of the same order as the size of the target region, which is typically many pixels in size. Thus, the sharp edges which are a visual cue to the presence of a minelike object can be lost. Secondly, background regions and non-minelike objects can register with significant values of r' , since the data sets from such regions may be well-correlated even if there is not large offset from the mean.

The image is filtered on a pixel-by-pixel basis. At each pixel, three quantities are evaluated: the background levels; the pseudo-correlation coefficient r' for the target region S centred on the pixel; and the differences of the pixel value in each data set from their respective background values. In order to discriminate against objects which do not stand out against the background, the pseudo-correlation coefficient r' is squared. Typically, pixels corresponding to minelike objects have somewhat higher values of r' than pixels corresponding to well-correlated areas of the background. Thus, squaring r' increases the separation between the values, tending to suppress background features. The filtered pixel value is the square of r' , multiplied by the greatest of the two differences from the mean and added to the desired overall filtered background level, say half the maximum pixel level.

The consideration of the difference from background at each pixel position has the result that sharp edges tend to stay sharp. It also means that large strong reflectors tend to be suppressed, since their size is large enough to raise the background level.

It has been mentioned above that it is not unusual to see evidence of distinct, across-track functional variations in the background signal level. These variations arise from mismatches between required and actual sonar TVG, and they tend to be most noticeable close to the centre of the sonar swath. In this area the mismatches occasionally result in rapid across-track variations in background signal level, with the pattern of the variations being similar for the two frequencies. Thus, considering equation (2.9), the value of r can be quite high. Because the target region is generally quite small, and assuming that there are no targets in the sonar record, even in the presence of a TVG mismatch the values of ξ_x and ξ_y will generally be quite small. The result, through the cosine term of equation (2.9), will be a relatively large value of r' with the net effect that the filtered image may appear to be 'banded' in the region of the midline. This problem can be corrected by making a simple modification to

equation (2.6), where the local background means for the square region S , \bar{X} and \bar{Y} , are replaced by a series of column-specific background means, $x_B[h_0, k]$ and $y_B[h_0, k]$, with h_0 representing the row coordinate of the centre of region S . An example of a suitable method for calculating x_B and y_B is included in the next section. The resulting, modified, pseudo-correlation coefficient then becomes

$$r' = \frac{\langle (x - x_B)(y - y_B) \rangle}{\left[\langle (x - x_B)^2 \rangle \langle (y - y_B)^2 \rangle \right]^{1/2}} \quad (2.10)$$

It is worth emphasising that, because the test region S is generally quite small and rapidly changing, local variations in background level are relatively unusual, the two pseudo-correlation coefficients (2.6) and (2.10) will generally behave almost identically. An additional advantage of using the modified pseudo-correlation coefficient is that a "circular buffer" approach can be adopted for the calculation of its components, allowing a worthwhile reduction in computation time.

2.2 Filter implementation

Let W be the "merged" sonogram, that is, the output of the filtration process when given input sonograms U and V , with pixels $w[h, k]$.

In the current filter implementation, the search or target region S is a "square", say

$$S_{h_0, k_0}(s) = \{[h, k]: h = h_0 - s, \dots, h_0 + s, k = k_0 - s, \dots, k_0 + s\} \quad (2.11)$$

where the size of the "square" is determined by the integer " s ", which is greater than or equal to one. Thus, the smallest squares are 3x3, 5x5, and so on. This shape was chosen for ease of implementation and because a minelike object is generally rectangular on the seabed, but with unknown orientation. The main control of filter behaviour is the parameter " s ".

As implemented, the filtration process required to produce the filtered value of the sonogram at location $[h_0, k_0]$, has three steps:

1. Calculate the "background levels" of the two sonograms across the width of the target square. For convenience, this is done at a prior stage in the calculation for *all* pixel positions along the row, using simple low-pass infinite impulse-response (IIR) (recursive) filters along each column of the sonograms, with the form:

$$\begin{aligned}
 x_B[h_0, k_i] &= \frac{1}{L} [x[h_0, k_i] + (L-1)x_B[h_0-1, k_i]] \\
 y_B[h_0, k_i] &= \frac{1}{L} [y[h_0, k_i] + (L-1)y_B[h_0-1, k_i]]
 \end{aligned}
 \tag{2.12}$$

where I is the range coordinate of the current column and L controls the "length" of the filter. Thus, the background region S' mentioned in section 2.1.2 is actually calculated as $2s+1$ weighted background levels, one for each column of pixels up to the current ping. The effective size of S' is determined by L . A series of "running averages" along the columns could also be used, but would incur a computational time penalty. If it is known that there will be no significant "TVG effects", then single background levels can be calculated for each pixel location. For instance averages may be calculated over a two dimensional region of which S is a subset. In this case the pseudo-correlation coefficient calculation described in the next paragraph would use equation (2.6).

2. Calculate the pseudo-correlation coefficient r' of the data set

$$\{x[h, k], y[h, k]: [h, k] \in S\} \tag{2.13}$$

using (2.10). The ordering of the elements of S is irrelevant in the calculation. If r' is negative, it may be set to zero, since anti-correlated regions are not of interest, however this has little effect on the resultant sonogram

3. The merged value of the sonogram, $w[h_0, k_0]$, is given by

$$w[h_0, k_0] = b + (r')^P \cdot \max \text{abs}(x[h_0, k_0] - x_B[h_0, k_0], y[h_0, k_0] - y_B[h_0, k_0]) \tag{2.14}$$

where b is an arbitrary, fixed, positive "background level", P is a positive exponent and "maxabs(p, q)" is "that one of p and q whose absolute value is the largest". The background level is added to allow the possibility of negative instances of the value selected by "maxabs".

IIR filters such as (2.12) and similar running summations are used throughout the filter implementation to reduce run-times and because they correspond well with the ping-by-ping nature of the recorded data.

2.3 Some examples

The examples which are included in this section were calculated with the "standard" pseudo-correlation coefficient (2.6). Figure 2 compares three views of the same section of sea bed containing a minelike object: at left is a section of 100 kHz sonogram, in the centre is the corresponding section of 500 kHz sonogram and at right is the merged

sonogram with the filter parameters from Section 2.3. The swath width is 100 metres, aligned up the page. The sonograms have been collected in relatively good conditions, however the lower contrast in the 500 kHz section and the higher resolution ensure that the mine is barely visible. In the merged image, the signatures of the minelike object and its attendant objects are clearly visible, but most of the smaller details on the bottom have been removed, one of the most useful features of this filter.

Figure 3 compares the values of the true correlation coefficient (2.3) (bottom left) and the pseudo-correlation coefficient (2.6) (bottom right) and the resultant merged images (top left and top right, respectively). The minelike object is the same one as is depicted in Figure 2. The images associated with the pseudo-correlation coefficient are clearly superior to those using the true correlation coefficient, in line with experience during the development of the filter. Note also that the "resolution" of both correlation coefficient images is limited by the use of the 7 by 7 moving square filter, so that they are much less sharp than the original images. The small squares in the correlation coefficient images correspond to isolated spikes occurring in both sonograms. When such a spike is inside the filter foreground region S , it dominates the summations and consequently all positions of S which contain the spike as one of their elements registers as a high coefficient value. Thus, the squares are at minimum 13 by 13 pixels in size.

Finally, Figure 4 compares original 100 and 500 kHz sonograms of a minelike object collected in conditions of high electrical interference with the merged sonogram of the same section of data. The merged sonogram is clearly superior to either of the originals in this case and the minelike object is more visible, primarily due to the reduction in the background noise level, rather than an increase in the absolute pixel level.

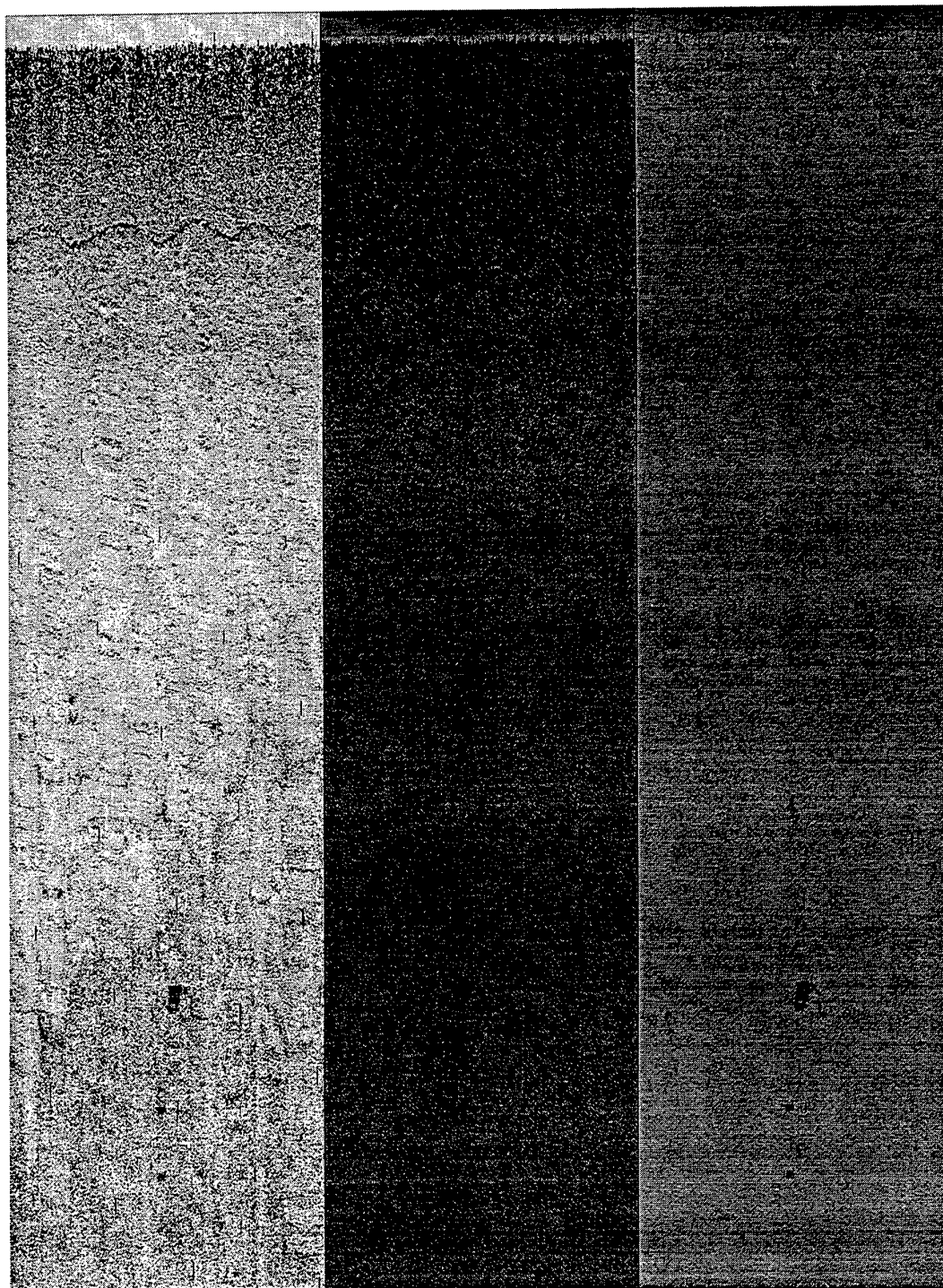


Figure 2. A comparison of the original corresponding 100 kHz (left) and 500 kHz (centre) sections of sonogram and the resultant merged sonogram (right). A reflective sonar target, a float and a multiple reflection of the float are visible in all three images, but most of the details of the background are absent from the merged image.

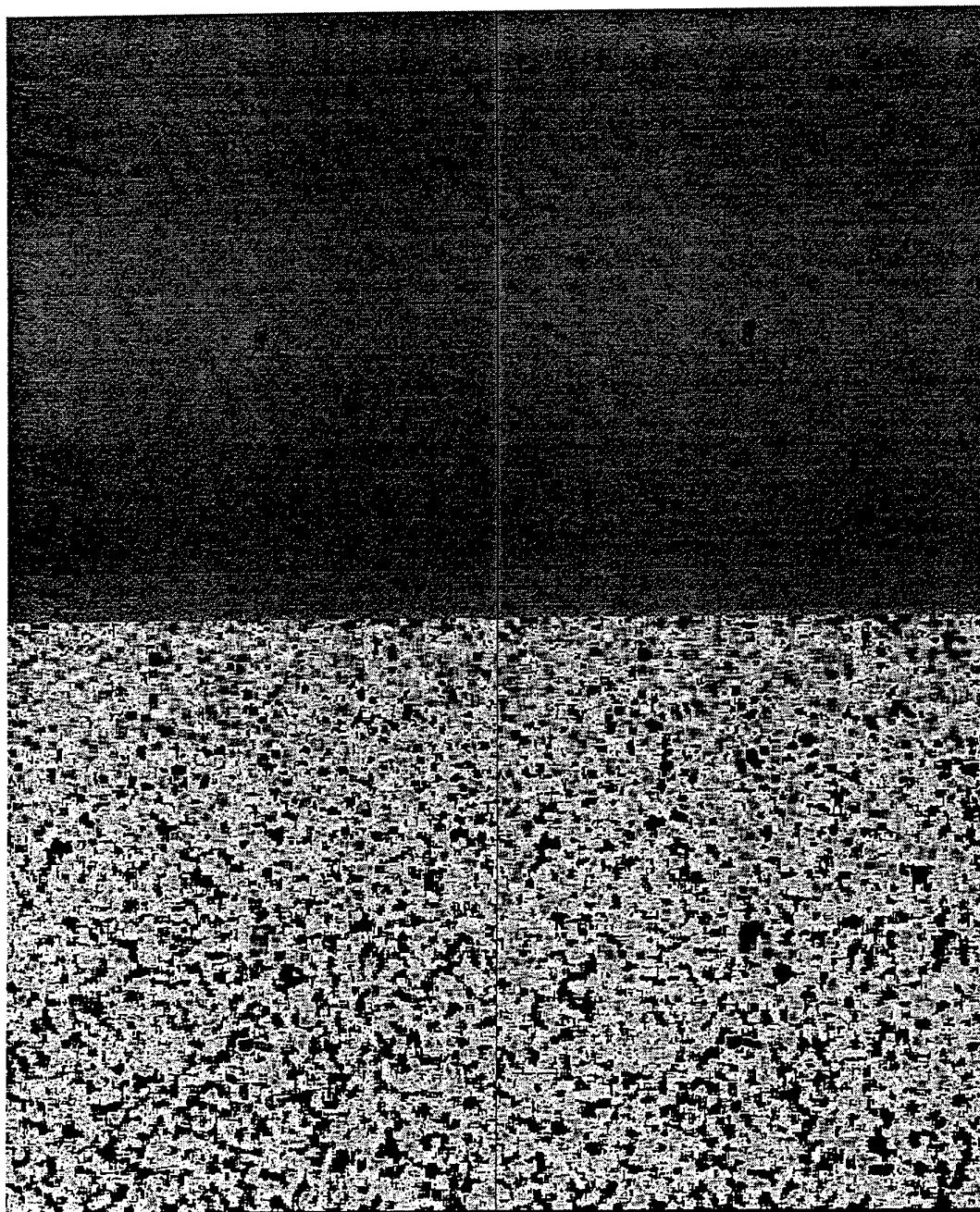


Figure 3. A comparison of the true correlation coefficient (bottom left) and pseudo-correlation coefficient (bottom right) in the region of the minelike object depicted in Figure 2. The corresponding merged images resulting from the use of each coefficient are shown at top left and top right respectively. The results from the pseudo-correlation coefficient are clearly superior to those from the true correlation coefficient. In this case strong correlation is represented black and weak correlation is represented white.

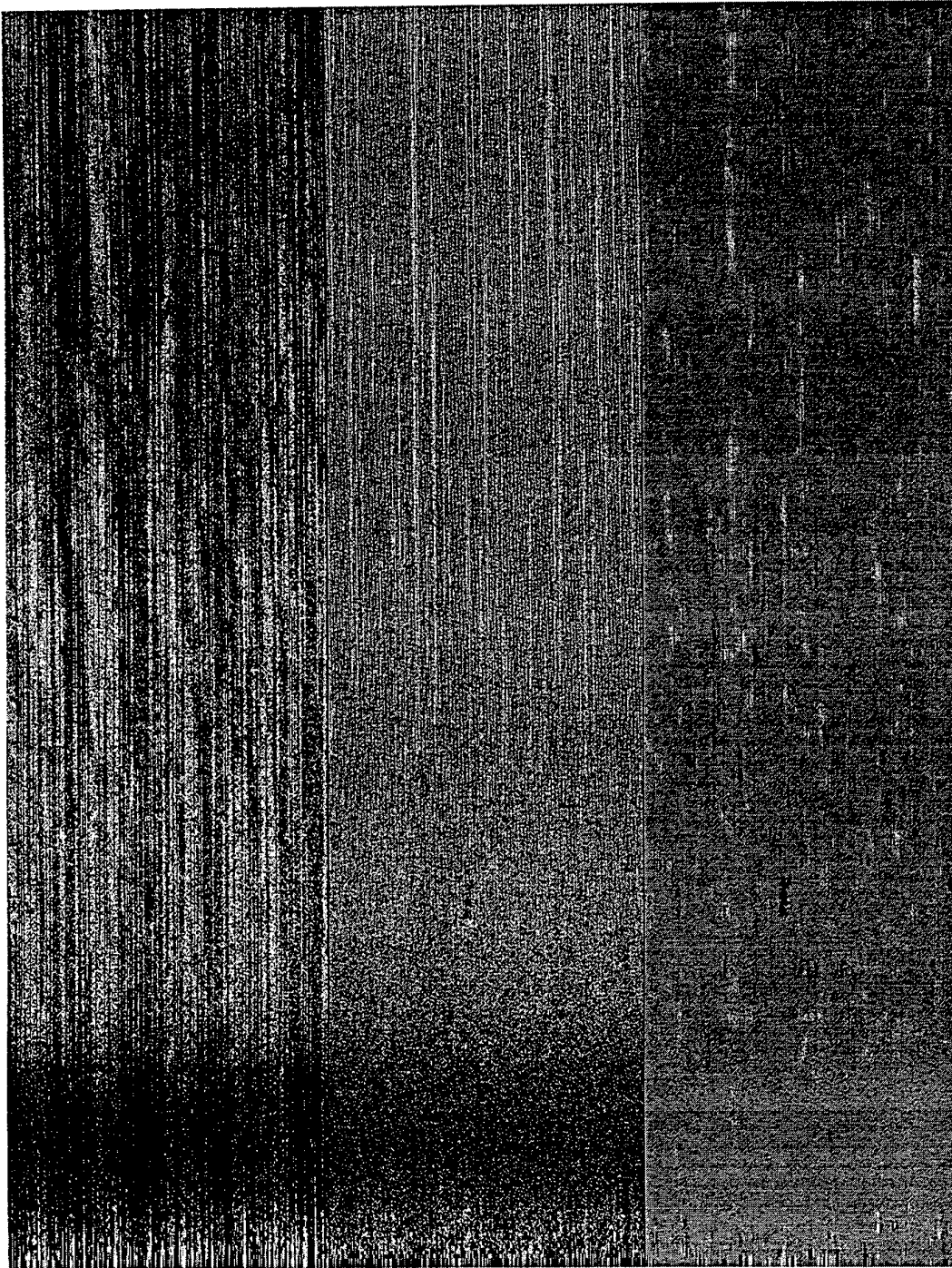


Figure 4. A comparison of corresponding sections of 100 kHz (left), 500 kHz (centre) and merged (right) sonograms for a particularly noisy data set. The 100 kHz sonogram is corrupted by streaking due to a faulty slip-ring connection and by surface reverberation. The 500 kHz sonogram is corrupted by smaller amounts of slip-ring noise. A minelike object is visible in the 500 kHz and merged images, but not in the 100 kHz image without processing.

2.4 Performance evaluation

Personal experience with the output of the merging filter tends to suggest that it is very robust, normally highlighting the presence of minelike objects at least as clearly and reliably to a human operator as inspection of both the 100 and 500 kHz sonograms. In many cases, minelike objects are easier to locate in the merged image than in either of the two original images, and this is particularly the case when one of the original images is strongly corrupted by surface reverberation.

In order to provide some kind of numerical measure of performance, a technique was devised which would estimate the "signal to noise" ratio of target (minelike) objects in the original and merged sonograms.

2.4.1 Definition of "local signal to noise ratio" for a target object in a sonogram

Consider a target object appearing in a sonogram. A "local signal to noise ratio" for the object should measure the amount by which the object differs from its immediate surroundings, since a human observer is able to discern such local variations in the image. The extent of a target object is taken to be limited to a rectangular region assigned to it by a human operator, using a mouse-controlled on-screen pointer to draw a box around the object, where the extent of the box is shown by a "rubber band" display. The size of the box is selected after comparing both original sonograms and the merged sonogram, and is set to contain the largest extent of the object in any of the three sonograms. The box contains all of the target pixels and may also contain other pixels, depending on the orientation of the object. The "background" is then taken to be two regions of equal width, but half the height of the target region, immediately above and below it. That is, if the target region T is defined by

$$T = \{[h, k]: h = h_0, \dots, h_1, k = k_0, \dots, k_1\} \quad (2.15)$$

then the background regions B_U and B_L (upper and lower) are given by

$$\begin{aligned} B_U &= \left\{ [h, k]: h = h_1 + 1, \dots, h_1 + \frac{h_1 - h_0}{2}, k = k_0, \dots, k_1 \right\} \\ B_L &= \left\{ [h, k]: h = h_0 - \frac{h_1 - h_0}{2}, \dots, h_0 - 1, k = k_0, \dots, k_1 \right\} \end{aligned} \quad (2.16)$$

The background region is thus $B = B_U + B_L$, which contains

$$N_B = (h_1 - h_0 + 1)(k_1 - k_0 + 1) \quad (2.17)$$

elements, the same number as the target region. Labelling the elements of B by i , the average background level and standard deviation are derived from

$$\begin{aligned}\langle\beta\rangle &= \frac{1}{N_B} \sum_{i=0}^{N_B-1} \beta_i \\ \sigma^2_{\beta} &= \frac{1}{N_B-1} \left(\sum_{i=0}^{N_B-1} \beta_i^2 - N_B \langle\beta\rangle^2 \right)\end{aligned}\tag{2.18}$$

where the β_i are the values of the appropriate sonogram (either of the originals u or v , or merged). The “ $n-1$ ” form of the standard deviation is used since the number of elements in B tends to be small.

The average background level and standard deviation are used to determine a threshold level for the elements of the target region. This is necessary, because usually the target object does not fill the entire rectangular target region and the target may have an associated acoustic shadow. The threshold value Θ_T is set two standard deviations higher than the mean background level, that is,

$$\Theta_T = \langle\beta\rangle + 2\sigma_{\beta}\tag{2.19}$$

with a separate value for each of the three sonograms. The target region of each of the three sonograms is now thresholded, and elements with values below the threshold are discarded. Let $\langle\tau\rangle$ be the mean of those elements of the target region whose signal levels are above the threshold. Then, the “local signal to noise ratio” of the target region in each sonogram is defined as the ratio of the signal power to the noise power, which is

$$\text{SNR} = \left(\frac{\langle\tau\rangle - \langle\beta\rangle}{\sigma_{\beta}} \right)^2.\tag{2.20}$$

2.4.2 Variability of the local signal to noise ratio

A concern with the use of the local SNR is that a judgement as to the limits of the target area is required from a human operator. The procedure used to date has been to compare the image of the target object in all three sonograms, then draw a box around it which is large enough to accommodate the largest observed extent. In order to measure the effects of this, some trials were done in which SNRs of the same minelike object were assessed repeatedly. They are given in Table 1 following. Each line in the table represents a separate trial and the values are grouped according to the box size condition which was being assessed. The first number in each column is the local SNR value, in dB and the second number in each column is the number of pixels in the box registering above the target threshold.

Table 1. Effects of box size on local SNR values for the same object

Box Size Condition	100 kHz	500 kHz	Merged
normal size box	13.8 (81)	9.0 (33)	22.3 (101)
	13.9 (76)	9.1 (30)	22.5 (94)
	13.4 (67)	8.5 (28)	20.6 (78)
sized to 500 kHz target	5.1 (32)	8.1 (24)	7.0 (22)
	2.6 (25)	7.8 (22)	none (0)
	5.1 (32)	8.1 (24)	7.0 (22)
sized to merge target	14.4 (93)	8.6 (32)	27.0 (128)
	13.9 (83)	8.3 (30)	24.5 (116)
	14.2 (90)	7.9 (29)	27.8 (128)
increasing sizes (too small to too big)	-4.2 (18)	3.7 (11)	none (0)
	3.2 (32)	8.3 (22)	6.7 (5)
	10.0 (58)	8.4 (28)	14.4 (67)
	13.8 (77)	8.9 (28)	23.2 (97)
	14.2 (83)	7.1 (25)	25.6 (110)
	13.9 (98)	7.6 (34)	26.2 (135)

Inspecting the table, it can be seen that the SNR values from different trials using approximately the same sized box usually give results stable to within 3 dB. It is also clear that sizing the box to the 500 kHz target alone is not advisable, and that the SNR values from the original sonograms are much less sensitive to box size than those from the merged sonogram, presumably because the backgrounds are noisier. Inspecting the number of pixels registering above threshold, it is clear that as the box size and hence the number of background pixels represented in the background standard deviation calculation increases, the target threshold decreases and more "target" pixels appear in the calculation. It is likely that box sizes far above the "normal" size do not lead to greatly different SNRs. In the evaluations following, the SNR values quoted are assumed to be valid within a margin of 3 dB.

3. Filter operation and performance

3.1 Analysis of Jervis Bay data sets up to 1995 with a fixed filter parameter

Local signal to noise ratios were calculated for minelike objects appearing in sonograms collected during various DSTO sea trials in Jervis Bay up to 1995, comprising 191 separate SNR evaluations for each of the three sonograms, too large a data set to include in this report. Some collected results are presented graphically.

The filter, as implemented for these evaluations, used the square of the correlation coefficient, and a 7 by 7 filter size was selected; that is, $P=2$ and $s=3$. In this case the background level setting b was not fixed, but was set equal the overall average of the background levels of the original sonograms, as output by the IIR filters (2.12).

Figure 5 shows the collected signal to noise ratios, in dB, calculated for all object types, plotted against across-track range. The SNR values for each sonogram are very scattered, but it is immediately obvious that targets in the merged sonogram score significantly higher than in either of the two "raw" sonograms. This agrees with the experience of the authors during numerous visual inspections of sonograms. Figure 6 shows the excesses of the SNRs from the merged sonograms over those from each of the two original sonograms for each particular target. Again, the values are almost all positive and some are very large. Detailed inspection of the data reveals that out of 191 separate mine and target identifications, there are 13 instances (6.8%) in which the signal to noise ratio of a target in the merged sonogram is exceeded by the signal to noise ratio in one or the other of the raw sonograms, and of these, 1 instance in which it is exceeded in both of the raw sonograms.

Almost invariably, the background standard deviation in the merged images is very low, that is, there is little noise power remaining in the sonogram. As a result of this, and recalling equation (2.20), very large (frequently over 30 dB) signal to noise ratios are a characteristic of many targets in the merged sonogram. The background in the merged sonograms is much less varied than in the original sonograms. This is because the pseudo-correlation coefficient in background areas is usually less than 1, and it is squared before being multiplied by the local difference between foreground and background regions, which is itself already a small number when the region considered does not contain a strong reflector. Note that visual inspection of the targets with high SNRs shows that they are indeed very obvious objects in the sonogram.

Given the large spread of values in all the SNR evaluations, it was decided to use the median score for each data set as the best indication of target visibility. Table 2 shows the median scores for the entire data set and for the various subsets corresponding to different classes and orientations of minelike objects. The spread of values is much lower than in Figures 5 and 6, which lends some credibility to the use of a median

score. The results reinforce the previous observations that the merging filter makes the minelike objects more readily visible than they were in the original images. It also seems to show that in general, the type of minelike object under consideration makes relatively little difference to the performance of the filter, or even to the visibility of the mine in the original sonograms.

Table 2. Median signal to noise ratios in dB for sea-trial data sets to 1995

Target type	100 kHz	500 kHz	merged
all possible	11.4	11.9	18.2
DSTO reflector 1	13.1	15.9	23.1
DSTO reflector 2	14.3	15.5	21.8
minelike 1a	14.4	10.6	23.1
minelike 1b	14.9	14.7	18.0
minelike 1c	9.2	12.1	18.2
minelike 1d	9.9	11.4	14.7
minelike 2	14.4	11.5	26.3
minelike 3a	14.6	11.4	22.8
minelike 3b	15.8	10.2	24.8
minelike 3c	8.9	11.4	17.4
minelike 3d	8.4	11.3	14.0
minelike 4	16.4	13.8	24.2

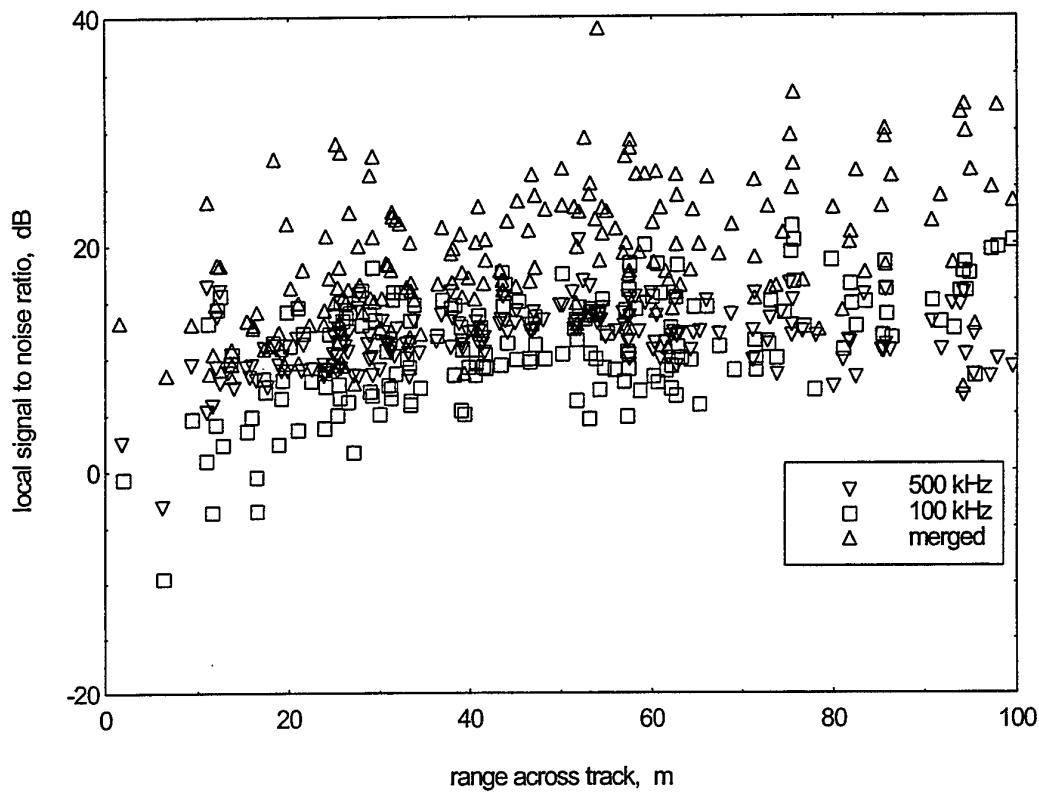


Figure 5. A comparison of local signal to noise ratios (SNRs) for minelike objects observed during DSTO sea trials recorded over a period of 3 years to 1995. In each instance, the image of the minelike object was inspected in all three sonograms, and the evaluation box for the local SNR was drawn so as to contain the largest of the images, the same box being used for all three images. The data were collected in a range of weather conditions and with many different types of minelike object, but SNR values for the merged sonograms are consistently larger than for either of the original sonograms, and this is true regardless of range.

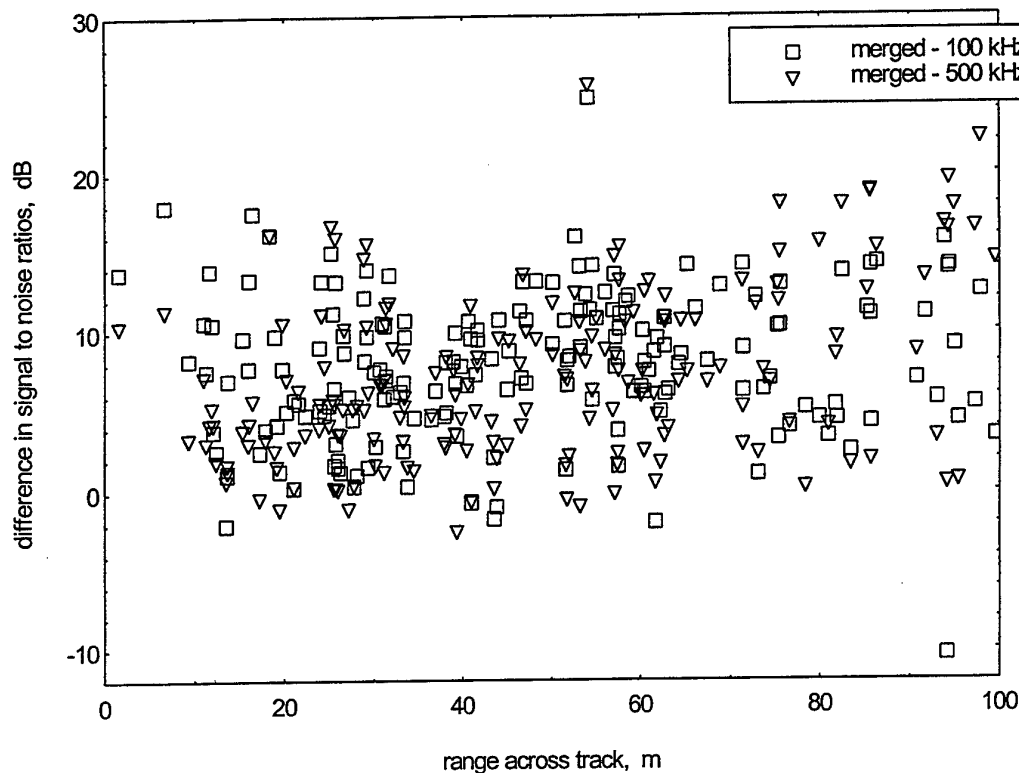


Figure 6. Differences between the local SNR values from the merged sonograms and the original 100 and 500 kHz sonograms for the 191 SNR evaluations from DSTO sea-trial data whose absolute SNR values are depicted in Figure 4. Most values are positive, indicating that the merged sonogram scored higher than the original sonogram.

3.2 Analysis of Jervis Bay 1996 data sets collected with different towfish configurations

After sea trials in Jervis Bay in 1996, further SNR data became available, this time involving only a two-week period of data collection. Data was analysed in the same manner as the earlier trials, except that in this case, the merge data background setting b was set to a fixed value of 100. In this case, the sonar configuration was subject to numerous changes during the trial, as different baffles were attached to the towfish in order to deliberately change the sound field entering the transducers from surface reverberation. After inspection of the data, it was not possible to distinguish a consistent effect on the SNR values due either to the different towfish configurations, nor to the different reflective targets. Table 3 shows the median SNR values from all data sets, and for each of the separate DSTO-constructed reflective targets. Note that there is as much variation between the median SNR values for targets 1 and 2, which are supposedly identical, as there is between the rest. The SNRs for the merged sonogram are however consistently higher than for the original sonograms.

Table 3. Median signal to noise ratios in dB from Jervis Bay sea trials, 1996

Data set	100 kHz	500 kHz	merged
all possible	10.9	8.3	18.8
DSTO reflector 1	8.6	7.5	16.4
DSTO reflector 2	10.2	8.7	18.1
DSTO reflector 3	11.4	8.4	19.9
DSTO reflector 4	10.9	8.1	19.5

Figure 7 shows the local SNR values calculated for the entire data set. The results are similar to those depicted in Figure 5 and again the merged sonogram scores consistently higher than the original sonograms. As the results for the merge data are comparable to those of earlier trials, the choice of background setting is clearly not critical to the performance of the filter.

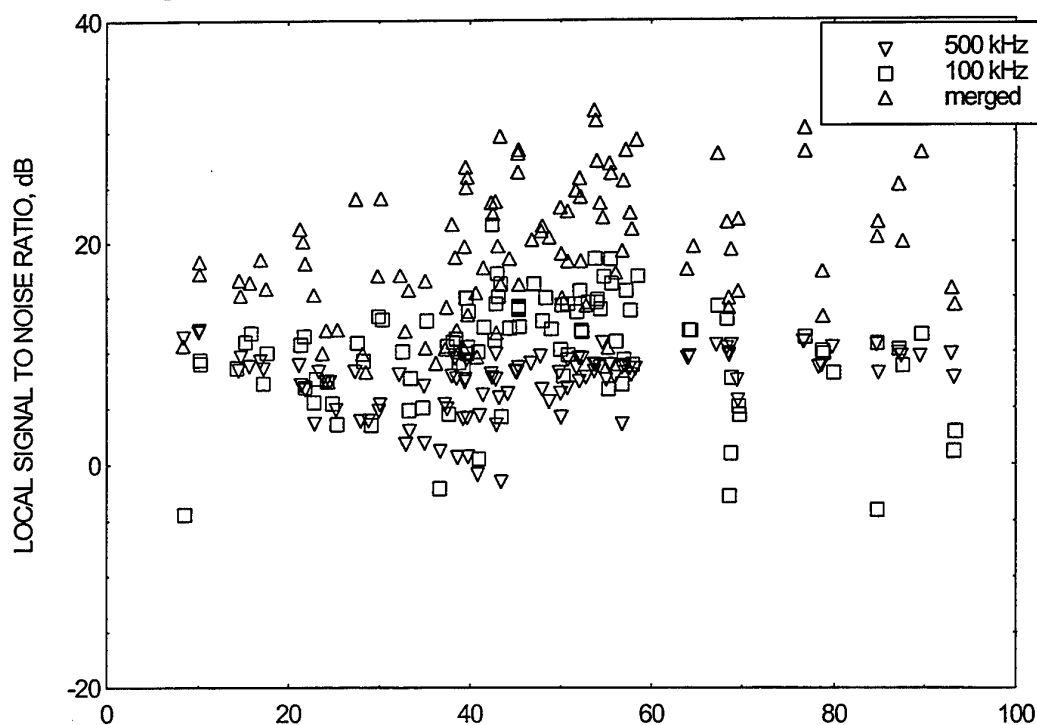


Figure 7. A comparison of local signal to noise ratios (SNRs) for minelike objects observed during DSTO sea trials recorded during a two-week period in 1996. The procedure for evaluating the SNRs was the same as is Figure 4. In contrast with Figure 4, the data in this figure were collected with different configurations of towfish attachments intended to block the path of sound, coming from above the towfish, into the transducers. Once again, however, the SNR values for the merged sonograms are consistently higher than those of either of the original sonograms, and this is true regardless of range.

3.3 Why the filter works

Considering Section 2.1.2, we see that the filter is controlled by the true correlation coefficient r of the 100 and 500 kHz sonograms, and by the "angles" ξ_x and ξ_y , which compare the differences of the target and background means with the respective target standard deviations. Minelike objects are not generally associated with the highest values of the true correlation coefficient registering in the image, which makes the value of r relatively unimportant in the calculation. In contrast, inspection of the target and background means and the target standard deviations corresponding to areas containing minelike objects shows that the differences between the means are mostly several times larger than the standard deviations. That is, it is usually true that at least one of ξ_x and ξ_y is large in the vicinity of a minelike object and they are rarely of opposite sign, so the pseudo-correlation coefficient expression is dominated by the product of sines. The value of the pseudo-correlation coefficient is usually close to 1 at minelike objects.

Some factors dependent on the nature of the Klein 590 dual-frequency sidescan sonar appear to assist the operation of the filter and allow it to be used at all ranges across the swath, which is taken here to be 100 m wide.

Firstly, at ranges above 50 m, the high rates of attenuation associated with the 500 kHz frequency mean that it is usually noise-limited, such that highly reflective objects register above the noise levels, but the less reflective common features of the bottom do not. Hence, r is often small everywhere except in the region of a target, since there is neither high true correlation, nor significant difference from the background.

Secondly, the beam of the 500 kHz transducer is less receptive to energy incident from above the horizontal than is the 100 kHz transducer. In data collected in shallow water, this means that surface reverberation (wave clutter) appearing in the 100 kHz sonogram is often much less evident or absent from the 500 kHz sonogram, which allows the filter to reduce the impact of surface reverberations on the merged sonogram. This is most apparent at large ranges across track.

Finally, considering the heavily streaked sonograms in Figure 4, the merging filter has been successful in this case because the noise levels in the sonograms were uncorrelated, so that r was small, and ξ_x and ξ_y were also not correlated, allowing the filter to successfully discriminate against the noise.

4. Conclusion

This report has demonstrated a filtration process which may be successfully and routinely used to reduce the impact of high noise levels in dual-frequency sonograms. It has been implemented as a PC-based computer program which is capable of real-time operation and is currently under assessment by personnel of the Royal Australian Navy. The filter is relatively simple to implement and offers considerable benefit in practice.

DISTRIBUTION LIST

"Merge" - A Filter for the Fusion of Dual-Frequency Sidescan Sonar Data

Roger A. Neill and Stuart D. Anstee

AUSTRALIA

1. DEFENCE ORGANISATION

a. Task Sponsor DGFD (Sea) - DDMWD

b. S&T Program

Chief Defence Scientist	} shared copy
FAS Science Policy	
AS Science Corporate Management	

Director General Science Policy Development
Counsellor Defence Science, London (Doc Data Sheet)
Counsellor Defence Science, Washington (Doc Data Sheet)
Scientific Adviser to MRDC Thailand (Doc Data Sheet)
Director General Scientific Advisers and Trials/Scientific Adviser Policy and
 Command (shared copy)
Navy Scientific Adviser
Scientific Adviser - Army (Doc Data Sheet and distribution list only)
Air Force Scientific Adviser
Director Trials

Aeronautical and Maritime Research Laboratory

Director
Chief of Maritime Operations Division
Alan Theobald
Brian Ferguson
Roger Neill
Stuart Anstee
John Best
Kurt Benke
David Hedger

Electronics and Surveillance Research Laboratory

Director (Doc Data Sheet only)
Tim Payne, LSOD

DSTO Library

Library Fishermens Bend
Library Maribyrnong
Library Salisbury (2 copies)
Australian Archives
Library, MOD, Pyrmont

c. Forces Executive

Director General Force Development (Land) (Doc Data Sheet only)

- d. **Navy**
 SO (Science), Director of Naval Warfare, Maritime Headquarters Annex,
 Garden Island, NSW 2000. (Doc Data Sheet only)
 COMAUSMINDIVFOR
 COMAUSMINDIVFOR-MWRSO
 Operations Manager, (MCD), MWSC
 DNW, Attn LCDR Peter Stokes
 Hydrographics Projects Officer, Hydrographics Projects Office
 MHD c/- Hydro Office, 8 Station St, Wollongong

 CMDR Paul Spencer, CO HMAS Moresby, Warship Section
 Perth Mail Exchange, WA 6758
- e. **Army**
 ABCA Office, G-1-34, Russell Offices, Canberra (4 copies)
- g. **S&I Program**
 Defence Intelligence Organisation
 Library, Defence Signals Directorate (Doc Data Sheet only)
- i. **B&M Program (libraries)**
 OIC TRS, Defence Regional Library, Canberra
 Officer in Charge, Document Exchange Centre (DEC), 1 copy
 *US Defence Technical Information Center, 2 copies
 *UK Defence Research Information Centre, 2 copies
 *Canada Defence Scientific Information Service, 1 copy
 *NZ Defence Information Centre, 1 copy
 National Library of Australia, 1 copy

2. UNIVERSITIES AND COLLEGES

Australian Defence Force Academy
 Library
 Head of Aerospace and Mechanical Engineering
 Deakin University, Serials Section (M list), Deakin University Library, Geelong, 3217
 Senior Librarian, Hargrave Library, Monash University
 Librarian, Flinders University

3. OTHER ORGANISATIONS

NASA (Canberra)
 AGPS

OUTSIDE AUSTRALIA

- 4. **ABSTRACTING AND INFORMATION ORGANISATIONS**
 INSPEC: Acquisitions Section Institution of Electrical Engineers
 Library, Chemical Abstracts Reference Service
 Engineering Societies Library, US
 Materials Information, Cambridge Scientific Abstracts, US
 Documents Librarian, The Center for Research Libraries, US

5. INFORMATION EXCHANGE AGREEMENT PARTNERS

Acquisitions Unit, Science Reference and Information Service, UK

Library - Exchange Desk, National Institute of Standards and Technology, US

TTCP GTP-13 National Leaders

USA- Doug Todoroff

Canada - Ron Kuwahara

NZ - Reinhard Holtkamp

UK - John Wickenden

SPARES (10 copies)

Total number of copies: 72

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA					
				1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)	
2. TITLE "Merge" - A Filter for the Fusion of Dual-Frequency Sidescan Sonar Data			3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION) Document (U) Title (U) Abstract (U)		
4. AUTHOR(S) Roger A. Neill and Stuart D. Anstee			5. CORPORATE AUTHOR Aeronautical and Maritime Research Laboratory PO Box 4331 Melbourne Vic 3001		
6a. DSTO NUMBER DSTO-TR-0511		6b. AR NUMBER AR-010-173		7. DOCUMENT DATE March 1997	
8. FILE NUMBER 510/207/0636		9. TASK NUMBER ADS 94/193		10. TASK SPONSOR DGFD (Sea)	
				11. NO. OF PAGES 23	
				12. NO. OF REFERENCES -	
13. DOWNGRADING/DELIMITING INSTRUCTIONS None			14. RELEASE AUTHORITY Chief, Maritime Operations Division		
15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT <i>Approved for public release</i> OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE CENTRE, DIS NETWORK OFFICE, DEPT OF DEFENCE, CAMPBELL PARK OFFICES, CANBERRA ACT 2600					
16. DELIBERATE ANNOUNCEMENT No Limitations					
17. CASUAL ANNOUNCEMENT Yes					
18. DEFTEST DESCRIPTORS Image processing; Data fusion; Route surveillance; Sidescan sonar; Target detection					
19. ABSTRACT A filtering and data fusion technique is described which uses the correlation between the two data streams of a dual-frequency sidescan sonar in order to discriminate against noise and preferentially enhance features of desired relative size. The filter is applied to data collected with a Klein 590 dual-frequency sidescan sonar with the parameters adjusted so as to enhance the display of minelike objects. Experience with the filter shows that it is highly effective in noisy conditions.					

TECHNICAL REPORT DSTO-TR-0511 AR-010-173 MARCH 1997



AERONAUTICAL AND MARITIME RESEARCH LABORATORY
GPO BOX 4331 MELBOURNE VICTORIA 3001
AUSTRALIA, TELEPHONE (03) 9626 7000